

### REMARKS

Claims 1-14 stands rejected under 35 USC §103(a) as being unpatentable over Byers, U.S. patent 6,219,645 in view of Gerson et al., U.S. patent 4,905,288.

Claims 15-20 stands rejected under 35 USC §103(a) as being unpatentable over Byers, U.S. patent 6,219,645 in view of Gerson et al., U.S. patent 4,905,288, and further in view of Namba et al., U.S. patent 5,884,249.

Claims 1, 9, 11, 14, and 15 have been amended to more clearly state the invention. Respective independent claims 1, 11, and 15, as amended, include subject matter of canceled dependent claims 8, 12, 13 and 20. Claims 8, 12, 13 and 20 have been canceled. Reconsideration and allowance of each of the pending claims 1-7, 9-11, and 14-19, as amended, is respectfully requested.

Byers, U.S. patent 6,219,645 discloses a method and system for simultaneously controlling a plurality of automatic speech recognition (ASR) systems within the working volume of a room, or for controlling multiple devices as a single unified ASR system. Multiple microphones feed a signal processing system which determines or estimates both a user's location and a user's orientation within a room as the user issues voice commands, and further determines the microphone providing the best signal. Diversity noise cancellation may be further applied to the microphone signals. Based on the apparent direction the user is facing, the system may then enable one of the ASR systems for voice command recognition and/or execution of voice commands. All of the signals from each of the microphones are directed to a signal processing circuit that compares the amplitude, frequency content, and phase of

the signals picked up from the various microphones, and decides the location of the user, and also estimates the user's head orientation. The signal processing circuit can select the microphone with the cleanest signal, e.g., using a digital signal processing technique to select the best of several spatially diverse sources. Sound takes a different amount of time to travel from the user to each of the microphones arrayed in the room. Each of the microphones receives corresponding signals, where the signal strength or amplitude, phases, arrival times, and frequencies are subject to the user's position and orientation within the room (the time the electrical signals take to get from the microphones to the signal processor is negligible). A delay number is related to the detected delay from the user to the microphone in question, and the microphone having the smallest delay number is initially presumed to be the closest microphone to the user.

Gerson et al., U.S. patent 4,905,288 discloses a method for data reduction in a speech recognition system and an arrangement for reducing a sequence of initial frames into a reduced set of representative frames by combining the initial frames into a plurality of representative frames, the combining process including generating a distortion measure associated with each representative frame and comparing each distortion measure to a distortion threshold. From these representative frames, a set of mutually exclusive frames is determined to minimize the number of representative frames, whereby each representative frame in the set represents a unique set of contiguous initial frames and has an associated distortion measure which does not exceed the distortion threshold. Initially, an untrained speaker-dependent

speech recognition system cannot recognize command words. The user manually prompts the device to begin the training procedure. Device controller 130 then directs switch 215 to enter the training mode. Device controller 130 then instructs speech synthesizer 240 to respond with the predefined phrase TRAINING VOCABULARY ONE, which is a "canned" response obtained from reply memory 260. The user then begins to build a command word vocabulary by uttering command words, such as STORE or RECALL, into microphone 205. The features of the utterance are first extracted by acoustic processor 110, and then applied to either word averager 220 or data reducer 230. Word averager 320 combines several utterances of the same word spoken by the user to provide a more reliable template. Speech recognizer 326, which performs the actual speech recognition comparison process, may use one of several speech recognition algorithms. The recognition algorithm of the present embodiment incorporates near-continuous speech recognition, dynamic time warping, energy normalization, and a Chebyshev distance metric to determine a template match. The recognizer control block 730 is used to coordinate the recognition process. Coordination includes endpoint detection (for isolated word recognition), tracking best accumulated distance scores of the word models, maintenance of link tables used to link words (for connected or continuous word recognition), special distance calculations which may be required by a specific recognition process and initializing the distance ram 734. The recognizer control may also buffer data from the acoustic processor. For each frame of input speech, the recognizer updates all active word templates in the template memory.

Namba et al., U.S. patent 5,884,249 discloses an input device which

includes plural inputting units, each accepting external information, including a unit for recognizing an input time of input information accepted via each of the plural inputting means; a unit for dividing or merging the input information into a predetermined primitive analysis unit, to obtain a recognition result; a unit for estimating the input time of a unit recognition result, by using an estimating method predetermined for each of the inputting units; and a unit for collecting some of the recognition results whose estimated input times are close to one another, and outputting the collected information as a semantic analysis unit. An input content is specified by collecting plural pieces of input information being signals input from plural inputting units. Column 1, lines 54-61 states: "The object of the present invention is to provide an inputting method for specifying an input content by collecting plural pieces of input information being signals input from plural inputting means. Another object of the present invention is to provide an input information managing method for specifying an input content by collecting input information being signals input from plural inputting means."

Independent claims 1, 11 and 15, as amended, respectively recite a method, computer program product, and apparatus for providing location-specific responses in an automated voice response system. Applicants respectfully submit that as amended, each of the independent claims 1, 11 and 15 is patentable over the references of record.

The method, computer program product, and apparatus for providing location-specific responses enables an automated, flexible and efficient voice response system including location-specific responses for controlling a plurality of different

devices.

As amended, independent claim 1 recites a method for providing location-specific responses in an automated voice response system. The method comprising the steps of: receiving a microphone signal from each of a plurality of microphones; identifying a spoken command utilizing speech recognition responsive to each said received microphone signal; storing a command start time and a command length for said identified spoken command and a channel number corresponding to one of said plurality of microphones utilizing said speech recognition; identifying a sound location vector responsive to each said identified spoken command utilizing said command start time, said command length for said identified spoken command and said channel number including performing digital signal analysis of said identified spoken command and applying said sound location vector to a central processor unit; and providing a response command based upon said sound location vector.

As amended, independent claim 11 recites a computer program product for providing location-specific responses in an automated voice response system including a processor, said computer program product including a plurality of computer executable instructions stored on a computer readable medium, wherein said instructions, when executed by a processor, cause the processor to perform the steps of: receiving a digitized audio signal from each of a plurality of microphones; utilizing speech recognition to identify a spoken command responsive to said received digitized microphone audio signal from each of a plurality of microphones including the steps of identifying said received microphone signal for a predetermined person and identifying

said spoken commands only from said identified predetermined person; identifying a sound location vector responsive to each identified spoken command by performing digital signal analysis for each identified spoken command utilizing a stored command start time  $T_0$ , and a command length  $T_c$  for said identified spoken command and a channel number of each identified one said plurality of microphones for each identified spoken command for identifying said sound location vector; and providing a response command based upon said sound location vector.

As amended, independent claim 15 recites apparatus for providing location-specific responses in an automated voice response system comprising: a plurality of microphones located within a defined environment for receiving a sound within said environment and each of said plurality of microphones providing a microphone signal; a respective speech recognition unit coupled to each one of said plurality of microphones; each said speech recognition unit for identifying spoken commands responsive to said microphone signal and for storing a command start time, a command length for said identified spoken command and a channel number corresponding to one of said plurality of microphones utilizing said speech recognition unit; a digital analysis unit coupled to said speech recognition unit for identifying a locational origin of said spoken command within said environment utilizing said command start time, said command length for said identified spoken command and said channel number; and applying said identified locational origin to a processor; and said processor for providing a response command based upon said identified locational origin of said spoken command within said environment; wherein said processor

selecting one of a plurality of predefined response commands utilizing said spoken command locational origin to provide said response command.

Applicant respectfully submits that the references of record do not disclose, nor suggest the method, computer program product, and apparatus for providing location-specific responses in an automated voice response system as taught by the present invention and recited in each of the independent claims 1, 11 and 15, as amended.

Applicant respectfully submits that each of the independent claims 1, 11 and 15, as amended, is patentable over all the references of record including Byers, Gerson et al. and Namba et al.

35 U.S.C. §103 requires that the invention as claimed be considered as a whole when considering whether the invention would have been obvious when it was made. In order to render an invention obvious by the combination of prior art references, the prior art must contain some reason, suggestion, or motivation to achieve the claimed invention, without using the Applicant's teaching as a guide for selecting and combining prior art elements. The mere fact that the prior art may be modified does not make the modification obvious unless the prior art suggested the desirability of the modification.

It is insufficient to establish obviousness that the separate elements of the invention existed in the prior art, absent some teaching or suggestion, in the prior art, to combine the elements. Applicant respectfully submits that the prior art descriptions of Byers, Gerson et al. and Namba et al. falls short of applicant's invention, and the

subject matter of the claimed invention as recited in each of the independent claims 1, 11 and 15, as amended, would not have been obvious to one of ordinary skill in the art in view of the references of record.

The prior art including Byers, Gerson et al. and Namba et al. do not disclose, nor suggest the recited method of independent claim 1, as amended, when considered as a whole, including the recited steps of identifying a spoken command and identifying a sound location vector, as set forth above. Only Applicant teaches these features of the method.

The prior art including Byers, Gerson et al. and Namba et al. do not disclose, nor suggest the recited computer program product of independent claim 11, as amended, when considered as a whole, including the recited steps of utilizing speech recognition to identify a spoken command and identifying a sound location vector responsive to each identified spoken command by performing digital signal analysis, as set forth above. Only Applicant teaches these features of the computer program product.

The prior art including Byers, Gerson et al. and Namba et al. do not disclose, nor suggest the recited apparatus of independent claim 15, as amended, when considered as a whole, including each said speech recognition unit for identifying spoken commands responsive to said microphone signal, and a digital analysis unit, as set forth above. Only Applicant teaches these features of the apparatus for providing location-specific responses in an automated voice response system.



Serial No. 09/941,112

Thus, each of the independent claims 1, 11 and 15, as amended, is patentable.

Dependent claims 2-7, 14 and 16-19 respectively depend from patentable independent claims 1, 11 and 15 and further define the invention. Thus, each of the dependent claims 2-7, 14 and 16-19 is patentable.

Applicants respectfully submit that the steps of adding a clock signal to each digitized microphone signal and the use of a clock adder as recited in representative dependent claims 2, 4, 17 and 18 is not shown nor suggested by the references of record. Applicant respectfully submits that the subject matter of dependent claim 6 is not shown nor suggested by the references of record.


Applicants have reviewed all the art of record, and respectfully submit that the claimed invention is patentable over all the art of record, including the references not relied upon by the Examiner for the rejection of the pending claims.

It is believed that the present application is now in condition for allowance and allowance of each of the pending claims 1-7, 9-11, and 14-19, as amended, is respectfully requested. Prompt and favorable reconsideration is respectfully requested.

If the Examiner upon considering this amendment should find that a telephone interview would be helpful in expediting allowance of the present application, the Examiner is respectfully urged to call the applicants' attorney at the number listed below.

Serial No. 09/941,112

Respectfully submitted,

By: 

Joan Pennington

Reg. No. 30,885

Telephone: (312) 670-0736